

Please replace the paragraph at page 8, line 19 through page 9, line 2 with the following paragraph:

B1  
The BLOCKS database (Steven Henikoff and Jorja G. Henikoff, "Automated assembly of protein blocks for database searching," *Nucleic Acids Research*, 19:23, pp. 6565-6572 (1991)) is an example of a database 17 of motifs. Emitof, and PRINTs are other such databases. These and other published databases may be used as the working predefined set/comparison database 17 in the present invention. Alternatively, it is possible to create a new motif database 17 from any protein database which has been labeled according to some parameter (e.g., structure). This is achieved by using multiple alignment software to find short multiply aligned ungapped sequences and then collecting statistics about these in a matrix. By creating a motif database 17 specific to the proteins of interest 11, more meaningful feature vectors 23 may be obtained since the motifs from a more general database may not occur in the proteins of interest.

Amendments to the specification are indicated in the attached "Marked Up Version of Amendments" (page i).

#### In the Claims

Please cancel claims 6 and 15.

Please add the following claims:

- B2
19. (New) A method for analyzing a subject protein sequence comprising the steps of:
- (a) providing a set of known biological fragments, the set being of a fixed number of said known biological fragments, each known biological fragment in the set having a respective representation;
  - (b) comparing the respective representation of each known biological fragment from the set to a subject protein sequence, for each known biological fragment said comparing including (i) counting the number of times the known biological fragment is found in the subject protein sequence and (ii) from said counted number of times,